

Studying the Z Boson with the ATLAS Detector at the LHC ¹

Q. Coc and P. Huth

Abstract: In this experiment we measure the mass and width of the Z -Boson using pre-selected data collected by the ATLAS detector at the LHC in 2012 along with Monte Carlo (MC) simulations of Standard Model processes [cite]. The comparison of both data sets, measurements and MC, reveal data points generated by secondary processes irrelevant for our purposes, which were consequently filtered out. By fitting the convolution of a Gaussian and a Breit-Wigner distribution to the filtered data we determine $m_Z = 90.629(5)$ GeV and $\Gamma = 3.258(21)$ GeV. Both values display a significant deviation from the values reported by the Particle Data Group ² of $\sigma_{M_Z} = 27.11$ and $\sigma_{M_Z} = 36.09$.

Als besondere Auswertung testiert: Datum, Unterschrift:

¹Experiment FP94 performed the 4th Nov. 2024. Supervisor: Ruiz, Miguel

²S. Navas et al. (Particle Data Group), Dec. 2024.[\[Online\]](#)

1 Introduction

The electroweak theory, unification of the electromagnetic and weak forces, was developed during the 70's by S. Glashow, A. Salam and S. Weinberg. It's lagrangian is symmetric under $SU(2)_L \times U_Y$ local gauge transformations and thus gives rise to 4 gauge bosons. After spontaneous symmetry breaking (SSB) the Goldstone-bosons corresponding to the weak interaction, W^+ , W^- and Z , acquire large masses while the boson mediating the electromagnetic, photon γ , remains massless. Five years after the publication of the electroweak theory, the W^\pm and Z boson were discovered experimentally.

The W and Z bosons were found at CERN. First, in the study of proton-antiproton collisions at the SPS and secondly in the analysis of electron-positron collision at the LEP. The latest precision measurements carried out by the ATLAS detector at the LHC.

The properties of the Z boson are mainly studied using the Drell-Yan process. It refers to the three-level Feynman diagram of a pair of quarks annihilating and forming a short lived virtual γ or Z bosons which subsequently decays into a pair of leptons: $q_c \bar{q}_c \rightarrow l \bar{l}$. From the properties of the final leptons it is possible to deduce information about the mediating boson. Since, such processes occur during highly energetic $p\bar{p}$ scattering, further processes and types of particles may contribute to the detected number of leptons.

Particle detection can be broken-down into three parts: reconstruction, calibration and identification. For this purpose the detector implements different components, namely: inner detector (ID), calorimeters (ECAL and HCAL), a muon spectrometer (MS) and a magnet system.

For the identification of electrons, the reconstruction algorithm uses information from the innermost layer (ID) and the calorimeter system. Due to the small ionization generated by muons, their reconstruction relies on the tracks from the ID and MS systems. These tracks are then fitted either by a e^- or a μ hypothesis to select candidates. The calibration is done by setting an appropriate energy scale with help of MC simulation. The candidates selected during this first stage are impure and can be corrected by a likelihood (LC) method, which quantifies the probability of an event to be an electron or muon, thus serving as an identification algorithm. Finally, the relevant particles can be selected by applying isolation requirements to further filter out leptons coming from undesired processes.

2 Procedure and Results

In the experiment FP91 we analyse properties of the Z boson using measurements taken in 2012 by the ATLAS experiment, which were published for educational use^{3 4}. The integrated luminosity is $\mathcal{L}_{int} = 1.0007(19) \text{ fb}^{-1}$ and the a center of mass energy $\sqrt{s} = 8 \text{ TeV}$. Additionally, we use MC simulations which were provided by our supervisor. Both of the datasets, measurements and MC, are tightly preselected.

We start the experiment by analysing different properties of the measured data. vx_{p-z} being the z-position of the primary vertex, depicts the z - position within ATLAS at which the primary vertex is located. With a mean positioning of -9.163 mm with respect to the center of the detector, collisions do show some variability in where they occur, possibly influenced by the nature of the utilized bundles and further calibrations to be done. The distribution also shows two distinct peaks as a result of further ongoing calibrations done during the data collection.

By plotting the number of leptons of each event in the measured data set we can see that most of the preselected events contain one lepton as final state. Such events are dominated by charged pion decays, e.g. $q \bar{q} \rightarrow \pi^- \rightarrow e^- \bar{\nu}_e$ and $q \bar{q} \rightarrow \pi^+ \rightarrow e^+ \nu_e$, revealing that most of the events are not relevant for studying the Z boson. These kind of events can be filtered out by imposing the number of leptons to be $n_l \geq 2$.

The lep_η variable contains the pseudorapidity which is, except for some areas, fairly constant. The gaps in the distribution are a result of the detector setup, like the extent which does not exceed

³"Review Studies for the ATLAS Open Data 8 TeV datasets, tools and activities", CERN, Dec. 2024.[\[Online\]](#)

⁴"Review Studies for the ATLAS Open Data Set", CERN, Dec. 2024.[\[Online\]](#).

$\eta \sim 2.5$. The azimuthal angle of the leptons is described by the $lep\text{-}\phi$ distribution and is constant throughout, due to the symmetry of the detector.

Lastly $lep\text{-}pt$ describes the transverse momentum of the measured leptons. Here, a rise around 25GeV can be observed, which is a result of the preselection process of the dataset, where each vertex has to have at least one lepton of $\geq 25\text{GeV}$. Nevertheless we do observe leptons with both greater and lesser transverse momentum. Leptons with higher momenta do naturally arise, but are less likely and thus also less numerous. The measured leptons with a lower transverse momenta are a result of decays that produced multiple leptons, of which at least one satisfied the preselection criteria.

For a further analysis of the measured data we calculate the invariant mass of the two leading leptons and compare it to the invariant mass of the $Z \rightarrow e^- e^+$ simulations, see 1. From theory, we would expect a resonance curve centred around the expected mass of the Z boson, $\sim 91.187\text{GeV}$. Our predictions are met by the invariant mass distribution of the MC simulations. While the mass distribution of the preselected data shows a predominant resonances peak around M_Z , three extra resonance curves are visible at $M < 50\text{GeV}$. These peaks correspond to bottom and charm resonances which likely decay into leptons. These events are irrelevant for our purposes.

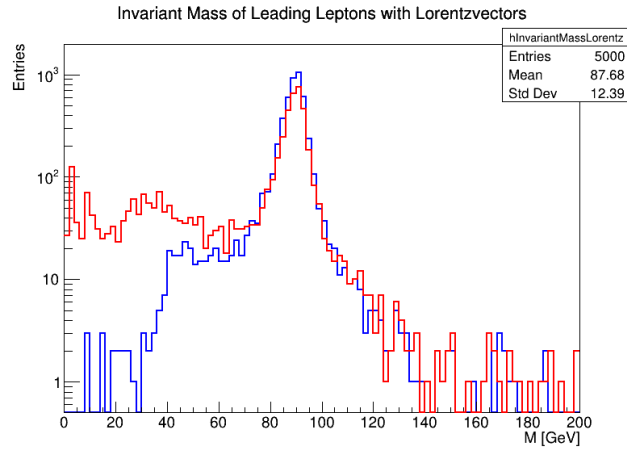


Figure 1: Leading lepton’s invariant mass of measured data (red) compared to the invariant mass of Monte Carlo simulation of $Z \rightarrow e^- e^+$ (blue). The invariant mass was computed using calculated ROOT’s `TLorentzVector` class.

After these two observations it is clear that the preselected data has to be filtered out to leave only the relevant Drell-Yan processes. For this purpose we impose isolation requirements to the measured data sets. To keep track of how many events are discarded by each applied criteria we plot a so called cut flow diagram in Figure 2. Additionally, we plot the invariant mass distribution of the filtered data. We do this for Z process with e , μ or τ pairs as final state. These plots are presented in Fig 3.

The bins on the cut-flow histogram represent applied isolation requirements. These constraints are applied consecutively from left to right. The y-axis shows the remaining number of events after the requirement on the x-axis was applied. Due to already harshly filtered data which is being used, no major change is generated by the first three conditions. For the same reason, there is no decrease in the number of events in the case of τ when Tight ID is applied. The number of events is unchanged between Lepton PDGID and Opposite Charges. For the case of τ , Trigger Fired was omitted, since the data lacks a `trigT` attribute. For the Isolation requirement we apply an isolation of 0.1.

From the cut flow histograms it is clear that the Lepton PDGID and Isolation requirements generate the biggest decrease in number of events. We conclude that cut on lepton-level constraints are more efficient than cut at event-level. Lepton-level cuts are specialized for the process to be analysed and thus are more efficient to filter out undesired events.

The corrected mass distributions from e and μ now show a prominent resonance peak around the expected mass of the Z boson. In contrast, the events filtered as τ show an asymmetric distribution which tends to $\sim 80\text{GeV}$. The reason for this behaviour is due to the properties of τ . Since

these particles are unstable, they don't reach the detector and decay into other particles. Hence, the filtered events contain the leptons coming from τ decays and don't give high quality information to study the Z boson. These kind of events could be identified by the p_T of the resulting neutrinos, $\nu_{e/\mu}$ and ν_{τ} , if these were able to be measured, which ATLAS is not capable of.

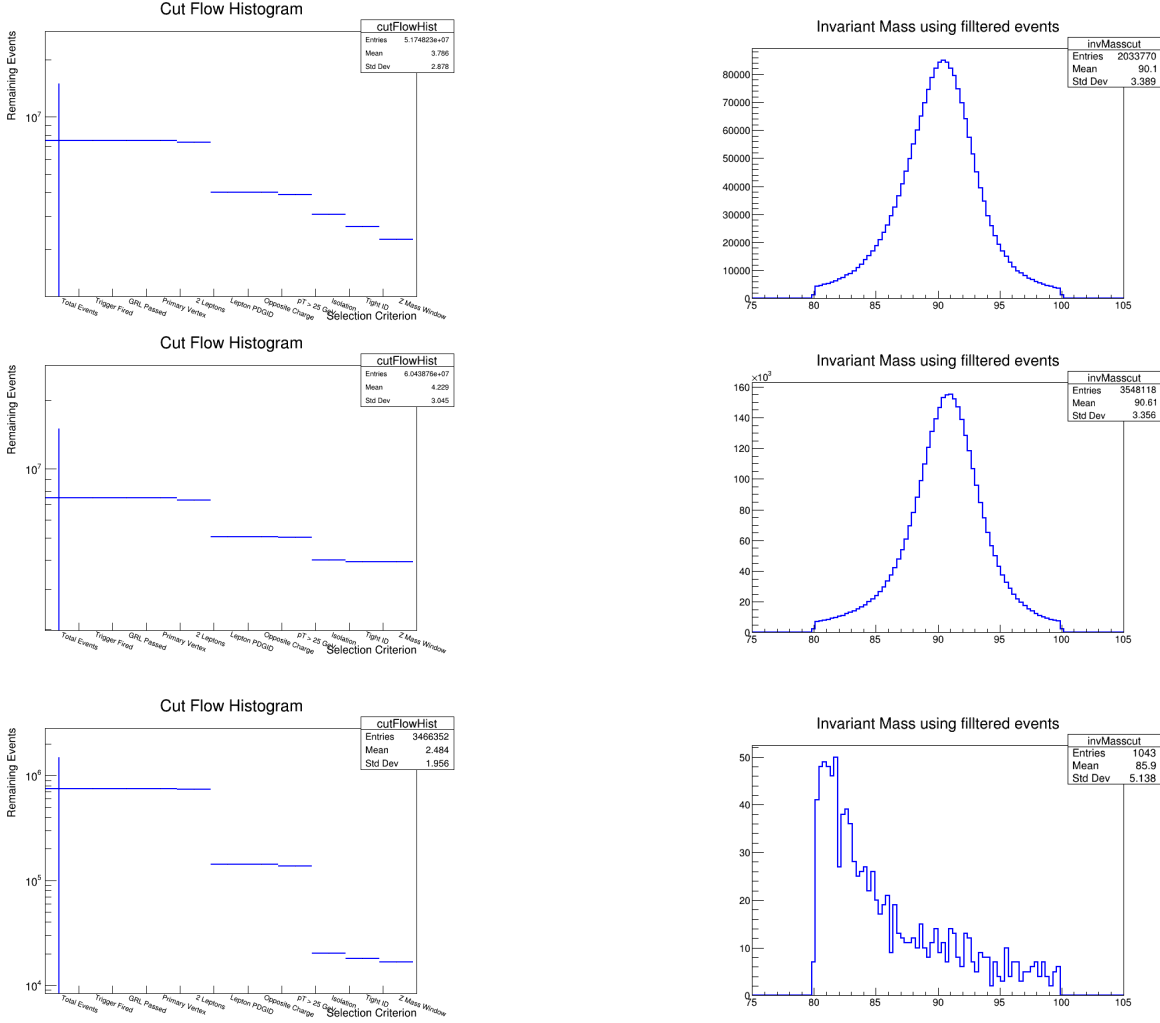


Figure 2: Cut flow from e , μ and τ events and invariant mass distribution of filtered data. Each row corresponds to the plots of each lepton: first row e , the second row μ and the third τ . The cut flow histogram is displayed on the first column and the invariant mass distribution on the left.

To study the properties of the Z boson we will only consider the e and μ pair events. For this reason we merge the filtered $Z \rightarrow e^-e^+$ and $Z \rightarrow \mu^-\mu^+$ events into a single histogram and compare it to the invariant mass distribution of the MC simulations, see right image in Fig. 3. The plot demonstrates the great agreement between the measured events and the MC simulations.

Next, we compare the fit of a Gauss distribution, a relativistic Breit-Wigner (BW) distribution and the resulting distribution from the convolution of the former ones. On one hand, due to the large number of data points in our data set and the nature of the measured events, the central limit theorem dictates that the distribution of the measurements tends to a normal distribution. On the other hand, from QFT we know that the mass distribution of an unstable particles in vacuum can accurately be described by a relativistic Breit-Wigner distribution. The left image in Fig. 3 makes clear that the BW renders a better fit compared to the gaussian fit. Nevertheless the χ^2/dof is rather high. The convolution of both distribution shows a clear advantage over the individual fits, since it take into consideration the statistical and physical nature of the measured data.

From the fitted convolution, we are able to determine the mass of the Z boson and its width. The

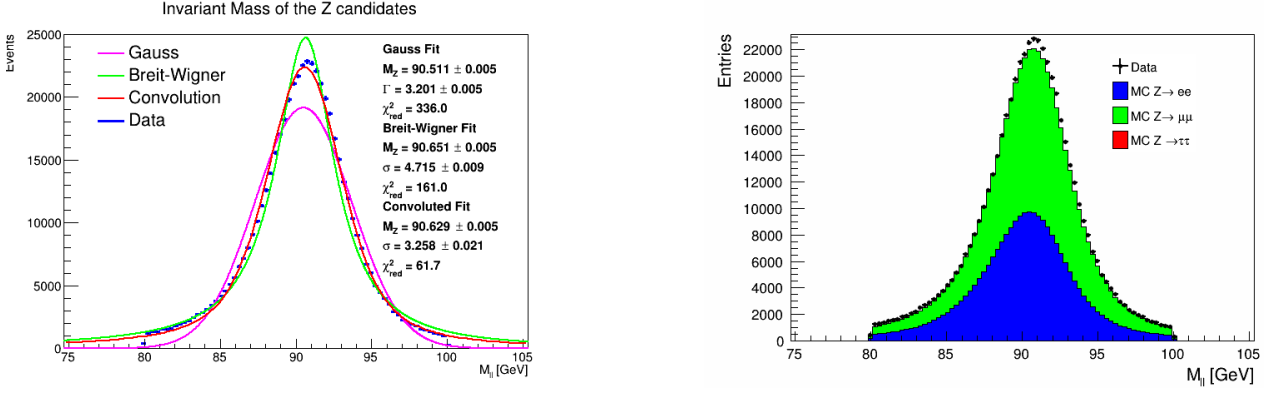


Figure 3: Right: Selected data (blue) and the fit of three distributions: (violet) Gauss-distribution, (green) Breit-Wigner Distribution and (red) convolution of a Gauss and Breit-Wigner distribution. Left: Merge of filtered electron and muon pair events (black cross) compared to MC simulations for $Z \rightarrow l\bar{l}$ processes.

experimentally found values are presented in Tab. 1 along with the values reported by the Particle Data Group ⁵. Even with the convolution yielding a better fit than any single distribution, we notice that the resulting mass and width for the Z boson deviate significantly when compared to the values reported by PDG. We will discuss possible error sources in the discussion.

Table 1: Comparison of measured quantities with values reported by the Particle Data Group (PDG).

	PDG	Fit parameter	Deviation
M_Z [GeV]	90.629(5)	91.1880(20)	27.11
Γ [GeV]	3.258(21)	2.4955(23)	36.09

As a final part of our analysis we use the tag-and-probe method to determine the efficiency of the different isolation requirements used for the selection of the data. Using the characteristic signature of $Z \rightarrow ee$ decays, the tag-and-probe method aims to produce a clean and unbiased sample to deduce said efficiency.

We introduce tight constraints upon one of the two electrons produced by the decay, in a similar fashion to the cut flow. The electron upon which these constraints are bestowed is the 'tag'. If it complies with the set restrictions, the second electrons, or the 'probe', will be further considered. Previous criteria on the invariant mass do still hold true and aim to further reduce background.

The efficiency is then obtained via the fraction of probe candidates that satisfy all criteria. The results of this analysis is depicted in Fig.4 against p_T .

3 Summary

In this experiment we examined a dataset provided by CERN from the ATLAS Detector, that was recorded in 2012 and attempted to calculate the Mass of the Z -Boson using the ROOT software framework, also provided by CERN.

We familiarized ourselves with the basics of ROOT's python frontend and functionalities like the TBrowse while exploring the plethora of available variables provided in the dataset. We specifically looked at the distributions of the numbers of leptons produced, z-Position of the primary vertex, transverse momentum, pseudorapidity and azimuthal angle observed in the collisions.

Limiting ourselves to two lepton events, we calculated the invariant mass using our own approach and using the provided TLorentzVectors, the results of which do match, but are easier to obtain via the

⁵S. Navas et al. (Particle Data Group), Dec. 2024.[\[Online\]](#)

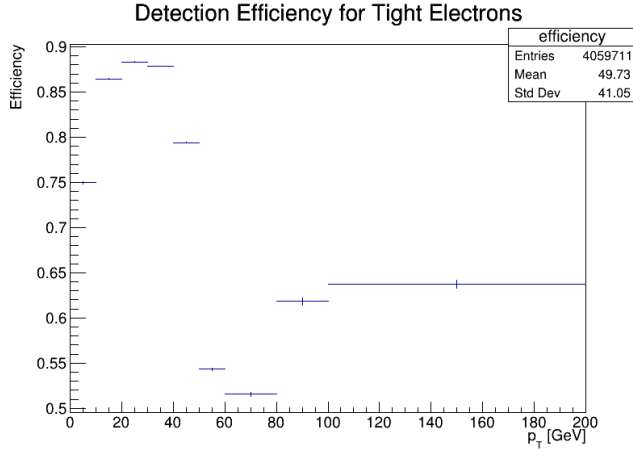


Figure 4: Detection Efficiency for tight Electrons obtained via the Tag-and-Probe method.

methods provided by ROOT. The distribution of the Invariant Mass was further compared to Monte Carlo simulations, revealing additional resonance curves below the expected peak of $\sim 91.187\text{GeV}$, corresponding to bottom and charm resonances which likely decayed into leptons.

To limit our data to the relevant Drell-Yan process we imposed isolation requirements upon the data like a Good run list, electron/muon trigger, restrictions on charge and so on. The effect of which can be observed in the cut flow histogram 2, isolating the relevant data greatly for e and μ events.

A convolution of a Breit-Wigner distribution and a Gauss distribution, was then fitted on the further selected data, yielding the final results of: $M_z = 92.1880\text{GeV}$ and $\Gamma = 2.4955\text{GeV}$, compared to the values provided by the Particle Data Group (PDG): $M_{Z,PDG} = 90.629\text{GeV}$ and $\Gamma_{PDG} = 3.258$, we observe significant deviations of: $\sigma_{M_z} = 27.11$ and $\sigma_\Gamma = 36.09$.

Lastly we had a brief examination of the efficiency using the tag and probe method. Implementing a similar routine to the cut flow specifically for the $Z \rightarrow ee$ decay, where the tag electron was subject to strict selection criteria, and if compliant, the complementary probe electron was measured. This yielded the Fig.4.

A discussion of the systematic errors, and deviations from expected results was undertaken in the critical discussion.

4 Critical discussion

Considering the sheer amount of effort that is put into the construction, maintenance and operation of the LHC, and the consideration of errors, addressing these is not an easy task. The preselected data might already have been subject to some corrections, that we can hardly estimate. For the purposes of our analysis, we saw that not all data was relevant, or could be used due to being the result of another processes, or producing only a certain number of leptons. Deviations due to the utilized software are unlikely, since it is the software provided and utilized by CERN. More directly the implementation of the cut flow, or the fit via a convolution of distributions could lead to some deviations. In our case, the fit model does not perfectly match the peak. One could modify the BW distribution and address the positioning and (a-)symmetry of the curve. Additionally the e^- could be missing some energy due to bremsstrahlung, for which one could include further corrections.

Lastly when examining vxp_z we saw two peaks in the distribution, which were caused by ongoing calibrations in early 2012, while the public data was recorded. Thus there might be some inconsistencies within this limited dataset when compared to the data accessible by the likes of PDG.